

From ranking to intransitive preference learning: rock-paper-scissors and beyond

Tapio Pahikkala¹, Willem Waegeman², Evgeni Tsivtsivadze³, Tapio Salakoski¹,
Bernard De Baets²

¹ Turku Centre for Computer Science (TUCS)
University of Turku, Department of Information Technology
Joukahaisenkatu 3-5 B, FIN-20520 Turku, Finland
`firstname.lastname@utu.fi`

² KERMIT, Department of Applied Mathematics, Biometrics and Process Control,
Ghent University, Coupure links 653, B-9000 Ghent, Belgium
`firstname.lastname@ugent.be`

³ Institute for Computing and Information Sciences, Radboud University Nijmegen,
Toernooiveld 1, 6525 ED Nijmegen, The Netherlands
`E.Tsivtsivadze@science.ru.nl`

Abstract. In different fields like decision making, psychology, game theory and biology, it has been observed that paired-comparison data like preference relations defined by humans and animals can be intransitive. The relations may resemble the well-known game of rock-paper-scissors. In the game, rock defeats scissors and scissors defeat paper, but rock loses to paper. Intransitive relations cannot be modelled with existing machine learning methods like ranking models, because these models exhibit strong transitivity properties. More specifically, in a stochastic context, where often the reciprocity property characterizes probabilistic relations such as choice probabilities, it has been formally shown that ranking models always satisfy the well-known strong stochastic transitivity property. Given this limitation of ranking models, we present a new kernel function that together with the regularized least-squares algorithm is capable of inferring intransitive reciprocal relations in problems where transitivity violations cannot be considered as noise. In this approach it is the kernel function that defines the transition from learning transitive to learning intransitive relations, and the Kronecker-product is introduced for representing the latter type of relations. In addition, we empirically demonstrate on two benchmark problems in game theory and biology that our algorithm outperforms methods not capable of learning intransitive reciprocal relations.

1 Introduction

We start with an introductory example in the field of sports games in order to describe the purpose of this paper. Let us assume that an online betting company for tennis games wants to build statistical models to predict the probability that a given tennis player will defeat his/her opponent in the next Grand Slam

competition. The company could be interested in building such models to maximize its profit when defining the amount of money that a client gets if he/she is able to predict the outcome of the game correctly. To this end, different types of data could be collected in order to construct the model, such as previous game outcomes, strong and weak points of players, current physical and mental conditions of players, etc. Yet, which type of machinery is required to obtain accurate predictions in this type of data mining problems? Firstly, as we will discuss in more detail below, we are for this example looking for an algorithm capable of predicting reciprocal relations from data, i.e., a relation between couples of players leading to a probability estimate of the outcome of a game. Secondly, we are also looking for a model that can predict intransitive relations, since commonly in sports games it turns out that game outcomes manifest cycles such as player A defeating player B, B defeating a third player C, and simultaneously C winning from A.

So, this paper in general considers learning problems where intransitive reciprocal relations need to be learned. As mathematical and statistical properties of human preference judgments, reciprocity and transitivity have been a subject of study for researchers in different fields like mathematical psychology, decision theory, social choice theory, and fuzzy modeling. Historically, this kind of research has been motivated by the quest for a rational characterization of human judgments, and to this end, transitivity is often assumed as a crucial property [1]. This property basically says that a preference of an object \mathbf{x}_i over another object \mathbf{x}_j and a similar preference of \mathbf{x}_j over a third object \mathbf{x}_k should always result in a preference of \mathbf{x}_i over \mathbf{x}_k , if preference judgments are made in a rational way. Nevertheless, it has been observed in several psychological experiments that human preference judgments often violate this transitivity property (see e.g. [2, 3]), especially in a context where preference judgments are considered as uncertain, resulting in non-crisp preference relations between objects.

To express uncertainty, we adopt a probabilistic framework in which it can be assumed that a preference relation defined on a space \mathcal{X} satisfies the reciprocity property.

Definition 1. *A function $Q : \mathcal{X}^2 \rightarrow [0, 1]$ is called a reciprocal relation if for any $(\mathbf{x}, \mathbf{x}') \in \mathcal{X}^2$ it holds that*

$$Q(\mathbf{x}, \mathbf{x}') + Q(\mathbf{x}', \mathbf{x}) = 1.$$

While taking into consideration this reciprocity property, [4] introduced several stochastic transitivity properties like weak, moderate and strong stochastic transitivity to characterize rational preference judgments in a probabilistic sense. We here recall the definition of weak stochastic transitivity.

Definition 2. *A reciprocal relation $Q : \mathcal{X}^2 \rightarrow [0, 1]$ is called weakly stochastically transitive if for any $(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) \in \mathcal{X}^3$ it holds that*

$$Q(\mathbf{x}_i, \mathbf{x}_j) \geq 1/2 \wedge Q(\mathbf{x}_j, \mathbf{x}_k) \geq 1/2 \Rightarrow Q(\mathbf{x}_i, \mathbf{x}_k) \geq 1/2. \quad (1)$$

This definition of transitivity for reciprocal relations naturally extends the basic definition of transitivity for crisp relations. Below, when we speak about intransitive relations, we specifically allude to relations violating weak stochastic transitivity. In addition, we will also utilize strong stochastic transitivity a few times in this paper. This stronger condition is defined as follows.

Definition 3. *A reciprocal relation $Q : \mathcal{X}^2 \rightarrow [0, 1]$ is called strongly stochastically transitive if for any $(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) \in \mathcal{X}^3$ it holds that*

$$Q(\mathbf{x}_i, \mathbf{x}_j) \geq 1/2 \wedge Q(\mathbf{x}_j, \mathbf{x}_k) \geq 1/2 \Rightarrow Q(\mathbf{x}_i, \mathbf{x}_k) \geq \max(Q(\mathbf{x}_i, \mathbf{x}_j), Q(\mathbf{x}_j, \mathbf{x}_k)).$$

Many other transitivity properties for reciprocal relations have been put forward in recent years, but these properties will not be discussed here. Moreover, many of these properties can be elegantly expressed in the cycle-transitivity framework. We refer to [5] for an overview of this framework and the various transitivity properties it covers.

The motivation for building intransitive reciprocal preference relations might be debatable in a traditional (decision-theoretic) context, but the existence of rational transitivity violations becomes more appealing when the notion of a reciprocal preference relation is defined in a broader sense, like in the introductory example, or generally as any binary relation satisfying the reciprocity property. For example, reciprocal relations in game theory violate weak stochastic transitivity, in situations where the best strategy of a player depends on the strategy of his/her opponent — see e.g. the well-known rock-scissors-paper game [6], dice games [7–9], and quantum games in physics [10]. Furthermore, in biology many examples of intransitive reciprocal relations have been encountered in competition between bacteria [11–15], and fungi [16], mating choice of lizards [17] and food choice of birds [18].

Generally speaking, we believe that enough examples exist to justify the need for models that can represent intransitive reciprocal relations. In this article we will address the topic of constructing such models based on any type of paired-comparison data. Basically, one can interpret these models as a mathematical representation of a reciprocal preference relation, having parameters that need to be statistically inferred. As a solution, we will extend an existing kernel-based ranking algorithm that has been proposed recently by some of the present authors [19]. This algorithm has been called RankRLS, as it optimizes a regularized least-squares objective function on paired-comparison data that is represented as a graph.

2 From transitive to intransitive preference models

In order to model preference judgments one can distinguish two main types of models in decision making [20, 21]:

1. *Scoring methods:* these methods typically construct a continuous function of the form $f : \mathcal{X} \rightarrow \mathbb{R}$ such that:

$$\mathbf{x} \succeq \mathbf{x}' \Leftrightarrow f(\mathbf{x}) \geq f(\mathbf{x}'),$$

which means that alternative \mathbf{x} is preferred to alternative \mathbf{x}' if the highest value was assigned to \mathbf{x} . In decision making, f is usually referred to as a utility function, while it is called a ranking function in machine learning.

2. *Pairwise preference models*: here the preference judgments are modeled by one (or more) valued relations $Q : \mathcal{X}^2 \rightarrow [0, 1]$ that express whether \mathbf{x} should be preferred over \mathbf{x}' . One can distinguish different kinds of relations such as crisp relations, fuzzy relations or reciprocal relations.

The former approach has been especially popular in machine learning for scalability reasons. The latter approach allows a flexible and interpretable description of preference judgments and has therefore been popular in decision theory and the fuzzy set community.

The semantics underlying reciprocal preference relations is often probabilistic: $Q(\mathbf{x}, \mathbf{x}')$ expresses the probability that object \mathbf{x} is preferred to \mathbf{x}' . One can in general construct such a reciprocal or probabilistic preference relation from a utility model in the following way:

$$Q(\mathbf{x}, \mathbf{x}') = g(f(\mathbf{x}), f(\mathbf{x}')), \quad (2)$$

with $g : \mathbb{R}^2 \rightarrow [0, 1]$ usually increasing in its first argument and decreasing in its second argument [22]. Models based on reciprocal preference relations have been applied in a machine learning context by, for example, [23].

The representability of reciprocal and fuzzy preference relations in terms of a single ranking or utility function has been extensively studied in domains like utility theory [24]. It has been shown that the notions of transitivity and ranking representability play a crucial role in this context.

Definition 4. *A reciprocal relation $Q : \mathcal{X}^2 \rightarrow [0, 1]$ is called weakly ranking representable if there exists a ranking function $f : \mathcal{X} \rightarrow \mathbb{R}$ such that for any $(\mathbf{x}, \mathbf{x}') \in \mathcal{X}^2$ it holds that*

$$Q(\mathbf{x}, \mathbf{x}') \leq \frac{1}{2} \Leftrightarrow f(\mathbf{x}) \leq f(\mathbf{x}').$$

Reciprocal preference relations for which this condition is satisfied have also been called weak utility models. [4] proved that a reciprocal preference relation is a weak utility model if and only if it satisfies weak stochastic transitivity, as defined by (1). As pointed out by [22], a weakly ranking representable reciprocal relation can be characterized in terms of (2) such that for any $(a, b) \in \mathbb{R}^2$ the function $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfies $g(a, b) > 1/2 \Leftrightarrow a > b$, $g(a, b) = \frac{1}{2} \Leftrightarrow a = b$. Analogous to weak ranking representability or weak utility models, one can define other conditions on the relationship between Q and f , leading to (stronger) transitivity conditions like moderate and strong stochastic transitivity. These properties are satisfied respectively by moderately and strongly ranking representable reciprocal preference relations. For such relations one imposes additional conditions on g , for example the following type of relations satisfies strong stochastic transitivity [4].

Definition 5. A reciprocal relation $Q : \mathcal{X}^2 \rightarrow [0, 1]$ is called *strongly ranking representable* if it can be written in the form of (2) with g given by

$$g(f(\mathbf{x}), f(\mathbf{x}')) = G(f(\mathbf{x}) - f(\mathbf{x}')), \quad (3)$$

where $G : \mathbb{R} \rightarrow [0, 1]$ is a cumulative distribution function satisfying $G(0) = \frac{1}{2}$.

In addition, other transitivity conditions and corresponding conditions on G have been defined, such as strict ranking representability. A further discussion on ranking representability is however beyond the scope of this paper.

3 Learning intransitive reciprocal relations

In this section we will show how intransitive reciprocal relations can be learned from data with kernel methods. During the last decade, a lot of interesting papers on preference learning have appeared in the machine learning community, see e.g. [25–28]. Many of these authors use kernel methods to design learning algorithms. The majority of them also considers utility approaches to represent the preferences. Only a few authors such as [29, 30] talk about pairwise preference relations, assuming weak stochastic transitivity so that an underlying ranking function exists.

We first explain the basic ideas behind kernel methods, followed by a discussion of a general framework for learning intransitive reciprocal relations. In this framework ranking can be seen as a special case, with a particular choice of the kernel function. To learn intransitive reciprocal relations, we then define a new type of kernel over pairs of data objects. Our analysis indicates that by using this kernel, we always learn relations that are reciprocal, but do not necessarily fulfill weak stochastic transitivity. This new kernel can be seen as a general concept that can be plugged into other kernel-based ranking methods as well, but in this paper we will illustrate its usefulness with the RLS algorithm. As this method optimizes a least-squares loss function, it is very suitable for learning reciprocal relations if the mean squared error measures the performance of the algorithm.

3.1 A brief introduction to kernels

This section is primarily based on [31, 32]. A better and much more detailed introduction to kernel methods can be found in these works. Given a not further specified input space \mathcal{E} that shows at this moment no correspondence with the space \mathcal{X} defined in the previous section, let us consider mappings of the following form: $\Phi : \mathcal{E} \rightarrow \mathcal{H}, \mathbf{e} \rightarrow \Phi(\mathbf{e})$. The function Φ represents a so-called feature mapping from \mathcal{E} to \mathcal{H} and \mathcal{H} is called the associated feature space. Initially, kernels were introduced to compute the dot-product $\langle \cdot, \cdot \rangle$ in this feature space efficiently. Such a compact representation of the dot-products in a certain feature space \mathcal{H} will in general be called a kernel with the notation $\langle \Phi(\mathbf{e}_1), \Phi(\mathbf{e}_2) \rangle = K(\mathbf{e}_1, \mathbf{e}_2)$.

Following the standard notations for kernel methods, we formulate our learning problem as the selection of a suitable function $h \in \mathcal{F}$, with \mathcal{F} a certain hypothesis space, in particular a kernel reproducing Hilbert space (RKHS). Hypotheses $h : \mathcal{E} \rightarrow \mathbb{R}$ are usually denoted as $h(\mathbf{e}) = \langle \mathbf{w}, \mathbf{e} \rangle$ with \mathbf{w} a vector of parameters that needs to be estimated based on training data. Let us denote a training dataset as a sequence

$$E = (\mathbf{e}_i, y_i)_{i=1}^N, \quad (4)$$

of input-label pairs, then we formally consider the following variational problem in which we select an appropriate hypothesis h from \mathcal{F} for training data E . Namely, we consider an algorithm

$$\mathcal{A}(E) = \operatorname{argmin}_{h \in \mathcal{F}} \frac{1}{N} \sum_{i=1}^N L(h(\mathbf{e}_i), y_i) + \lambda \|h\|_{\mathcal{F}}^2 \quad (5)$$

with L a given loss function and $\lambda > 0$ a regularization parameter. The first term measures the performance of a candidate hypothesis on the training data and the second term, called the regularizer, measures the complexity of the hypothesis with the RKHS norm. In our framework below, a least-squares loss $L(h(\mathbf{e}), y) = (h(\mathbf{e}) - y)^2$ is optimized in (5). Optimizing this loss function instead of the more conventional hinge loss has the advantage that the solution can be found by simply solving a system of linear equations. Due to lack of space we do not describe in details the mathematical properties and advantages of this approach compared to more traditional algorithms, but more details can be found for example in [33].

According to the representer theorem [31], any minimizer $h \in \mathcal{F}$ of (5) admits a dual representation of the following form:

$$h(\mathbf{e}) = \sum_{i=1}^N a_i K(\mathbf{e}, \mathbf{e}_i) = \langle \Phi(\mathbf{e}), \mathbf{w} \rangle,$$

where $a_i \in \mathbb{R}$, K is the kernel function associated with the RKHS mentioned above, Φ is the feature mapping corresponding to K , and $\mathbf{w} = \sum_{i=1}^N a_i \Phi(\mathbf{e}_i)$.

3.2 Learning reciprocal relations

We will use the above framework in order to learn intransitive reciprocal relations. To this end, we associate in a preference learning setting with each input a couple of data objects, i.e. $\mathbf{e}_i = (\mathbf{x}_i, \mathbf{x}'_i)$, where $\mathbf{x}_i, \mathbf{x}'_i \in \mathcal{X}$ and \mathcal{X} can be any set. Consequently, we have an i.i.d. dataset $E = (\mathbf{x}_i, \mathbf{x}'_i, y_i)_{i=1}^N$ so that for each couple in the training dataset a label is known. These labels will represent reciprocal relations observed on training data, but rescaled to the interval $[-1, 1]$. This means that the following correspondence holds

$$y = 2Q(\mathbf{x}, \mathbf{x}') - 1, \quad \forall (\mathbf{x}, \mathbf{x}') \in \mathcal{X}^2.$$

Such a conversion is primarily made for ease of implementation. This implies that we will minimize the regularized squared error so that a model of type $h : \mathcal{X}^2 \rightarrow \mathbb{R}$ is obtained. For the squared loss we can simply choose a function $G(a)$, whose value would be 0, $(a + 1)/2$, and 1, in case $a < -1$, $-1 \leq a \leq 1$, and $a > 1$, respectively, so that $[0, 1]$ -valued relations are predicted as $Q(\mathbf{x}, \mathbf{x}') = G(h(\mathbf{x}, \mathbf{x}'))$. To guarantee that reciprocal relations are learned, let us suggest the following type of feature mapping:

$$\Phi(\mathbf{e}_i) = \Phi(\mathbf{x}_i, \mathbf{x}'_i) = \Psi(\mathbf{x}_i, \mathbf{x}'_i) - \Psi(\mathbf{x}'_i, \mathbf{x}_i),$$

where Φ is just the same feature mapping as before but now written in terms of couples and Ψ is a new (not further specified) feature mapping from \mathcal{X}^2 to a feature space. As shown below, this construction will result in a reciprocal representation of the corresponding $[0, 1]$ -valued relation. By means of the representer theorem, the above model can be rewritten in terms of kernels, such that two different kernels pop up, one for Φ and one for Ψ . Both kernels express a similarity measure between two couples of objects and the following relationship holds:

$$\begin{aligned} K^\Phi(\mathbf{e}_i, \mathbf{e}_j) &= K^\Phi(\mathbf{x}_i, \mathbf{x}'_i, \mathbf{x}_j, \mathbf{x}'_j) \\ &= \langle \Psi(\mathbf{x}_i, \mathbf{x}'_i) - \Psi(\mathbf{x}'_i, \mathbf{x}_i), \Psi(\mathbf{x}_j, \mathbf{x}'_j) - \Psi(\mathbf{x}'_j, \mathbf{x}_j) \rangle \\ &= \langle \Psi(\mathbf{x}_i, \mathbf{x}'_i), \Psi(\mathbf{x}_j, \mathbf{x}'_j) \rangle + \langle \Psi(\mathbf{x}'_i, \mathbf{x}_i), \Psi(\mathbf{x}'_j, \mathbf{x}_j) \rangle \\ &\quad - \langle \Psi(\mathbf{x}_i, \mathbf{x}'_i), \Psi(\mathbf{x}'_j, \mathbf{x}_j) \rangle - \langle \Psi(\mathbf{x}'_i, \mathbf{x}_i), \Psi(\mathbf{x}_j, \mathbf{x}'_j) \rangle \\ &= K^\Psi(\mathbf{x}_i, \mathbf{x}'_i, \mathbf{x}_j, \mathbf{x}'_j) + K^\Psi(\mathbf{x}'_i, \mathbf{x}_i, \mathbf{x}'_j, \mathbf{x}_j) \\ &\quad - K^\Psi(\mathbf{x}'_i, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}'_j) - K^\Psi(\mathbf{x}_i, \mathbf{x}'_i, \mathbf{x}'_j, \mathbf{x}_j). \end{aligned}$$

Using this notation, the prediction function given by the representer theorem can be expressed as:

$$h(\mathbf{x}, \mathbf{x}') = \langle \mathbf{w}, \Psi(\mathbf{x}, \mathbf{x}') - \Psi(\mathbf{x}', \mathbf{x}) \rangle = \sum_{i=1}^N a_i K^\Phi(\mathbf{x}_i, \mathbf{x}'_i, \mathbf{x}, \mathbf{x}').$$

For this prediction function, we can easily show that it forms the basis of a reciprocal relation.

Proposition 1. *Let $G : \mathbb{R} \rightarrow [0, 1]$ be a cumulative distribution function satisfying $G(0) = 0.5$ and $G(-a) = 1 - G(a)$, then the function $Q : \mathcal{X}^2 \rightarrow [0, 1]$ defined by*

$$Q(\mathbf{x}, \mathbf{x}') = G(h(\mathbf{x}, \mathbf{x}')), \quad (6)$$

with $h : \mathcal{X}^2 \rightarrow \mathbb{R}$ given by (6), is a reciprocal relation.

3.3 Ranking: learning transitive reciprocal relations

Using the above notation, utility or ranking functions are usually written as

$$f(\mathbf{x}) = \langle \mathbf{w}, \phi(\mathbf{x}) \rangle.$$

They can be elegantly expressed in our framework by defining a specific feature mapping and corresponding kernel function.

Proposition 2. *If K^Ψ corresponds to the transitive kernel K_T^Ψ defined by*

$$K_T^\Psi(\mathbf{x}_i, \mathbf{x}'_i, \mathbf{x}_j, \mathbf{x}'_j) = K^\phi(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle,$$

with K^ϕ any two-dimensional kernel function on \mathcal{X}^2 , whose value depends only on the arguments \mathbf{x}_i and \mathbf{x}_j and their feature representations $\phi(\mathbf{x}_i)$ and $\phi(\mathbf{x}_j)$, then the reciprocal relation $Q : \mathcal{X}^2 \rightarrow [0, 1]$ given by (6) is strongly stochastically transitive.

For this choice of K^Ψ , our framework is reduced to a popular type of kernel function that has been introduced by [25]. The insight of the proposition is that the use of this kernel is equivalent to constructing a ranking for the individual inputs. This ranking function is in the dual representation given by:

$$f(\mathbf{x}) = \langle \mathbf{w}, \phi(\mathbf{x}) \rangle = \sum_{i=1}^N \alpha_i (K^\phi(\mathbf{x}_i, \mathbf{x}) - K^\phi(\mathbf{x}'_i, \mathbf{x})).$$

As explained in Section 2, ranking results in a reciprocal relations that satisfies the weak stochastic transitivity property. Due to the above proposition, we can even claim that the resulting reciprocal relation satisfies strong stochastic transitivity. Different ranking methods are obtained with different loss functions, such as RankSVM [34] for the hinge loss and RankRLS [19] for the least-squares loss.

3.4 Learning intransitive reciprocal relations

Since the above choice for Ψ forms the core of all kernel-based ranking methods, these methods cannot generate intransitive relations, i.e. relations violating weak stochastic transitivity. In order to derive a model capable of violating weak stochastic transitivity, we introduce the following feature mapping Ψ_I for couples of objects:

$$\Psi_I(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x}) \otimes \phi(\mathbf{x}'),$$

where $\phi(\mathbf{x})$ is again the feature representation of the individual object \mathbf{x} and \otimes denotes the Kronecker-product of matrices (see e.g. [35]). Kernel functions induced by this type of feature maps have also been considered under the name tensor product kernels (see e.g. [36]) and the Kronecker product has also been used to construct kernels based of linear feature transformation (see e.g. [37]).

In the following, we use the following property of the Kronecker product:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD),$$

where $A \in \mathbb{R}^{a \times b}$, $B \in \mathbb{R}^{c \times d}$, $C \in \mathbb{R}^{b \times e}$, and $D \in \mathbb{R}^{d \times f}$. The Kronecker-product establishes joint feature representations Φ_I and Ψ_I that depend on both arguments of Φ and Ψ . Instead of ignoring the second argument of Φ and Ψ , we now

represent all pairwise interactions between individual features of the two data objects in the joint feature representation. Using the notation K_I^Ψ , this leads to the following expression:

$$\begin{aligned} K_I^\Psi(\mathbf{x}_i, \mathbf{x}'_i, \mathbf{x}_j, \mathbf{x}'_j) &= \langle \phi(\mathbf{x}_i) \otimes \phi(\mathbf{x}'_i), \phi(\mathbf{x}_j) \otimes \phi(\mathbf{x}'_j) \rangle \\ &= \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle \otimes \langle \phi(\mathbf{x}'_i), \phi(\mathbf{x}'_j) \rangle \\ &= K^\phi(\mathbf{x}_i, \mathbf{x}_j) K^\phi(\mathbf{x}'_i, \mathbf{x}'_j), \end{aligned}$$

with again K^ϕ any kernel function defined over \mathcal{X}^2 . As a result, using the Kronecker-product as feature mapping basically leads to a very simple kernel in the dual representation, consisting of just a regular product between two traditional kernels K^ϕ . Remark that K^ϕ can be any existing kernel, such as the linear kernel, the RBF-kernel, etc. As a result of the above construction, the kernel function K^Φ becomes:

$$K_I^\Phi(\mathbf{x}_i, \mathbf{x}'_i, \mathbf{x}_j, \mathbf{x}'_j) = 2K^\phi(\mathbf{x}_i, \mathbf{x}_j)K^\phi(\mathbf{x}'_i, \mathbf{x}'_j) - 2K^\phi(\mathbf{x}'_i, \mathbf{x}_j)K^\phi(\mathbf{x}_i, \mathbf{x}'_j).$$

We further refer to K_I^Φ as the intransitive kernel.

Indeed, in the above extension of the ranking framework, two different kernels K^Ψ and K^Φ must be specified by the data analyst, while the third kernel K^Φ is defined by the choice for K^Ψ . On the one hand, the choice for K^Ψ (and hence K^Φ) determines whether the model is allowed to violate weak stochastic transitivity. On the other hand, the kernel function K^ϕ acts as the traditional similarity measure on \mathcal{X} , resulting in a linear, polynomial, radial basis function or any other representation of the data.

We now present a result indicating that the intransitive kernel K_I^Φ can be used to learn arbitrary reciprocal preference relations provided that the feature representation ϕ of the individual objects is powerful enough.

Proposition 3. *Let E be a training dataset of type (4), L be a loss function, and $\mathcal{F}_R : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ be the set of all hypotheses inducing a reciprocal relation on \mathcal{X} . Moreover, let*

$$h^* = \operatorname{argmin}_{h \in \mathcal{F}_R} \sum_{i=1}^N L(y_i, h(\mathbf{x}_i, \mathbf{x}'_i)) \quad (7)$$

be the set of hypotheses inducing a reciprocal relation on \mathcal{X} that have minimal empirical loss on E . Further, let

$$\begin{aligned} h(\mathbf{x}, \mathbf{x}') &= \sum_{i=1}^N \alpha_i K_I^\Phi(\mathbf{x}_i, \mathbf{x}'_i, \mathbf{x}, \mathbf{x}') \\ &= \sum_{i=1}^N \alpha_i 2 (K^\phi(\mathbf{x}_i, \mathbf{x}) K^\phi(\mathbf{x}'_i, \mathbf{x}') - 2K^\phi(\mathbf{x}'_i, \mathbf{x}) K^\phi(\mathbf{x}_i, \mathbf{x}')) \end{aligned} \quad (8)$$

be the set of hypotheses we can construct using the intransitive kernel K_I^Φ and a given feature representation ϕ of a base kernel K^ϕ .

There exists such a feature representation ϕ and such coefficients $a_{i=1}^N$ that the corresponding hypothesis (8) is one of the minimizers of (7).

The proof of the proposition is based on first calculating the minimal empirical loss one can obtain with a hypothesis inducing a reciprocal relation and then providing an example of a feature map ϕ that can be used to achieve this error. The complete proof is presented in [38]. The above result indicates that this type of model is flexible enough to obtain as low empirical error on training data as it is possible to get while maintaining the reciprocity property, and hence it can also learn intransitive reciprocal relations.

4 Experiments

4.1 Rock-Paper-Scissors

In order to test our approach, we consider a semi-synthetic benchmark problems in game theory, a domain in which intransitive reciprocal relations between players is often observed. In such a context, a pure strategy provides a complete description of how a player will play a game. In particular, it determines the move a player will make for any situation (s)he could face. A player’s strategy set is the set of pure strategies available to that player. A mixed strategy is an assignment of a probability to each pure strategy. This allows for a player to randomly select a pure strategy. Since probabilities are continuous, there are infinite mixed strategies available to a player, even if the strategy set is finite.

We consider learning the reciprocal relation of the probability that one player wins from another in the well-known rock-paper-scissors game. To test the performance of the learning algorithm in such a nonlinear task, we generated the following synthetic data. First, we generate 100 individual objects for training and 100 for testing. The data objects are three-dimensional vectors representing players of the rock-paper-scissors game. The three attributes of the players are the probabilities that the player will choose ‘rock’, ‘paper’, or ‘scissors’, respectively. The probability $P(r \mid \mathbf{x})$ of player \mathbf{x} choosing rock is determined by $P(r \mid \mathbf{x}) = \exp(wu)/z$, where u is a random number between 0 and 1, w is a steepness parameter, and z is a normalization constant ensuring that the three probabilities sum up to one. The probabilities for ‘paper’ and ‘scissors’ are determined analogously. By varying the width w of the exponential function, we can generate players tending to favor one of the three choices over the others or to play each choice almost equally likely.

We generate 1000 player couples for training by randomly selecting the first and the second player from the set of training players. Each couple represents a game of rock-paper-scissors and the outcome of this game can be considered as stochastic in nature, because the strategy of a player is chosen in accordance with the probabilities of picking a particular fixed strategy from that player’s set of mixed strategies. For example, when a fixed rock player plays against a mixed strategy player that plays scissors with probability 0.8 and paper with probability 0.2, then we have a higher chance of observing a game outcome for

which the fixed rock player wins from the second player. Yet, the same couple of players with different outcomes can simultaneously occur in the training data. During training and testing, the outcome of a game is -1 , 0 , or 1 depending on whether the first player loses the game, the game ends in a tie, or the first player wins the game, respectively. We use the game outcomes as the labels of the training couples.

For testing purposes, we use each possible couple of test players once, that is, we have a test set of 10000 games. However, instead of using the outcome of a single simulated game as label, we assign for each test couple the element of the reciprocal relation that corresponds to the probability that the first player wins:

$$Q(\mathbf{x}, \mathbf{x}') = P(p | \mathbf{x})P(r | \mathbf{x}') + \frac{1}{2}P(p | \mathbf{x})P(p | \mathbf{x}') + P(r | \mathbf{x})P(s | \mathbf{x}') \\ + \frac{1}{2}P(r | \mathbf{x})P(r | \mathbf{x}') + P(s | \mathbf{x})P(p | \mathbf{x}') + \frac{1}{2}P(s | \mathbf{x})P(s | \mathbf{x}').$$

The task is to learn to predict this reciprocal relation. The algorithm estimates the relation by rescaling the predicted outputs that lie in the interval $[-1, 1]$, as discussed above.

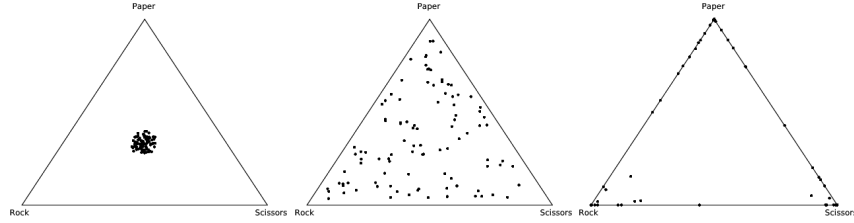


Fig. 1. Illustration of the players in the three data sets generated using the values 1 (top left), 10 (top right), and 100 (bottom) for the parameter w .

	$w = 1$	$w = 10$	$w = 100$
I	0.000209	0.000445	0.000076
II	0.000162	0.006804	0.131972
III	0.000001	0.006454	0.125460

Table 1. Mean-squared error obtained with three different algorithms: regularized least-squares with the kernel K_I^Φ (I), regularized least-squares with the kernel K_T^Φ (II) and a naïve approach consisting of always predicting $1/2$ (III).

We conduct experiments with three data sets generated using the values 1, 10, and 100 for the parameter w . These parameterizations are illustrated in Figure 1. The value $w = 1$ corresponds to the situation where each player tends

to play ‘rock’, ‘paper’, or ‘scissors’ almost equally likely, that is, the players are concentrated in the center of the triangle in the figure. For $w = 100$ the players always tend to play only their favorite item, that is, the players’ strategies are concentrated near the three corners of the triangle. Finally, $w = 10$ corresponds to a setting between these two extremes.

The results are presented in Table 1. We report the mean squared-error obtained by regularized least-squares in a transitive and intransitive setting, respectively by specifying the kernels K_T^ϕ and K_I^ϕ . For K^ϕ a simple linear kernel is chosen in both cases. We also compare these two approaches with a naive heuristic consisting of always predicting 1/2 (a tie). This heuristic can be interpreted as quite optimal for $w = 1$, because in that case all players are located in the center of the triangle. This explains why neither the transitive nor the intransitive regularized least-squares algorithm can outperform this naive approach when $w = 1$. We conclude that there is not much to learn in this case. For the other two values of w , the situation is different, with the regularized least-squares algorithm with the intransitive kernel performing substantially better than the naive approach, while the performance with the transitive kernel being close to that of the naive one. Unsurprisingly, learning the intransitive reciprocal relations is more difficult when the probabilities of the players are close to the uniform distribution ($w = 10$) than in case the players tend to always play their favorite strategy ($w = 100$). Especially in this last case, regularized least-squares with an intransitive kernel performs substantially better than its transitive counterpart. This supports the claim that our approach works well in practice, when the reciprocal relation to be learned indeed violates weak stochastic transitivity. The stronger this violation, the more the advantage of an intransitive kernel will become visible.

4.2 Theoretical Biology

Inspired by the simulations made by [39], we consider the following setting. Suppose we have a number of competing species, each of them having two features. Namely, a species \mathbf{x} has a strong point denoted by $s(\mathbf{x})$ and a weak point denoted by $w(\mathbf{x})$, and the values of both features are between 0 and 1. Then, for a couple of individuals, say $(\mathbf{x}, \mathbf{x}')$, we define a label y , whose value equals 1 if \mathbf{x} dominates \mathbf{x}' and -1 in the opposite case. The dominance is determined by the following formula:

$$y = \text{sign}(u(s(\mathbf{x}'), w(\mathbf{x})) - u(s(\mathbf{x}), w(\mathbf{x}')))) \quad (9)$$

where sign is the signum function and

$$u(a, b) = \min(|a - b|, 1 - |a - b|). \quad (10)$$

We observe that the species \mathbf{x} dominates \mathbf{x}' if and only if the strong point $s(\mathbf{x})$ of \mathbf{x} is closer to the weak point $w(\mathbf{x}')$ of \mathbf{x}' than $s(\mathbf{x}')$ is to $w(\mathbf{x})$, the closeness being defined by (10).

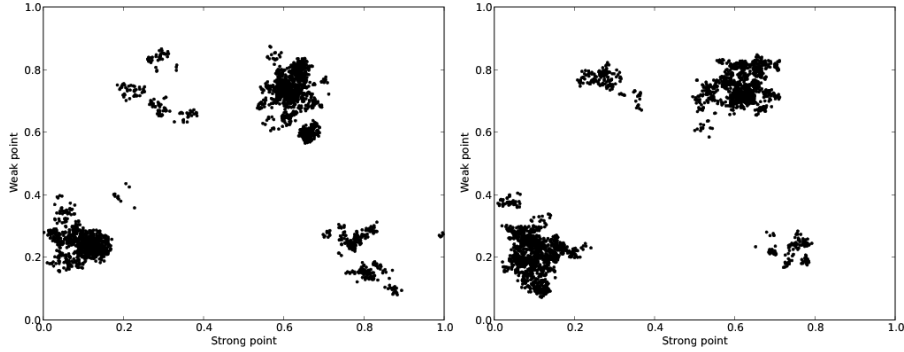


Fig. 2. The set of 2,500 species after 900,000 (left) and 1,000,000 (right) confrontations.

We set up an experiment in which we randomly generate an initial population of 2,500 species so that their strong and weak points have been drawn from a uniform distribution between 0 and 1. Then, we select randomly two species \mathbf{x} and \mathbf{x}' from the population for which we compute a label y with (9). In the confrontation of these two species, we say that x is the winner and \mathbf{x}' is the loser if $y = 1$ and vice versa if $y = -1$. After the confrontation, the winner replaces the loser with its own descendant $\hat{\mathbf{x}}$. The strong and weak points of the descendant are obtained from the strong and weak points of the winner by shifting them by a small amounts whose sizes are drawn from a normal distribution having zero mean and standard deviation 0.005.

Unlike in the experiments done by [39], we adopt an approach in which we do not consider any local neighborhood of the species, that is, the two confronting species are randomly selected from the current population of 2,500 species. In addition, for each confrontation of two species, there is always a winner and a loser, while this was the case in the experiments of [39] only if the value of (10) for $s(\mathbf{x})$ and $w(\mathbf{x}')$ was smaller than a certain threshold. A next couple was randomly selected in case of the value being larger than the threshold. Finally, our closeness function (10) differs from the one used by [39] so that the strong and the weak points are cyclic in the sense that values 0 and 1 can be considered to be equal. We adopted the cyclic property of the weak and strong points in order to eliminate the special case of the values being close to 0 and 1.

We perform altogether 1,000,000 subsequent confrontations of two species. In the beginning, there are no clusters, since the strong and weak points of the species are uniformly distributed. However, the species start to form small clusters after a couple of tens of thousands of confrontations and large clusters when a couple of hundreds of thousands of confrontations has passed. We sample our training and test sets from the 100,000 last confrontations, since at this point the simulation has already formed quite stable clusters. Namely, we randomly sample without replacement 1,000 couples for a training set and 10,000 for a

test set. The clusters formed after 900,000 and 1,000,000 confrontations are depicted in Figure 2.

I	0.849900
II	0.615200

Table 2. Classification accuracy obtained with two different algorithms: regularized least-squares with the kernel K_I^Φ (I) and regularized least-squares with the kernel K_T^Φ (II).

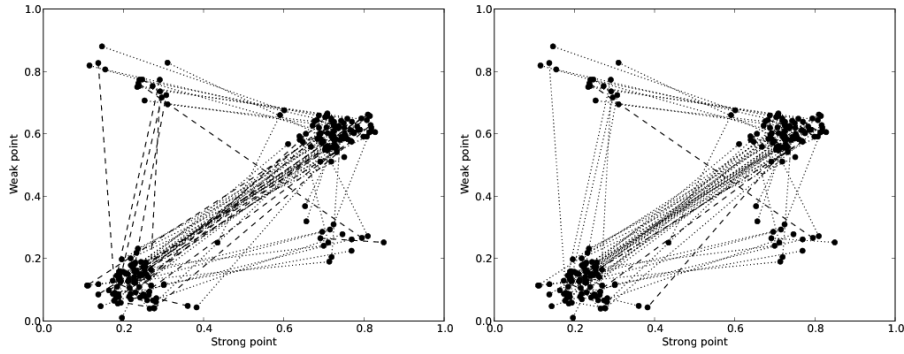


Fig. 3. Illustration of 100 randomly selected test couples. Left: the dotted lines denote the 69 couples classified correctly and the dashed lines denote the 31 incorrectly classified ones using RLS with the transitive kernel. Right: the dotted lines denote the 89 couples classified correctly and the dashed lines denote the 11 incorrectly classified ones using RLS with the intransitive kernel.

We train two RLS classifiers with the training set of 1000 confrontations and use them for predicting the outcomes of the unseen 10000 confrontations in the test set. The first classifier uses a transitive kernel K_T^Φ and the second one an intransitive kernel K_I^Φ . The base kernel K^ϕ is chosen to be the Gaussian radial basis function kernel for both the cases, that is, $K^\phi(\mathbf{x}, \mathbf{x}') = e^{-\gamma((s(\mathbf{x})-s(\mathbf{x}'))^2 + (w(\mathbf{x})-w(\mathbf{x}'))^2)}$. The value of the regularization parameter and the width γ of the Gaussian kernel are selected with a grid search and cross-validation performed on the training set. The classification accuracies for both classifiers are listed in Table 2. Moreover, a random sample of 100 test couples and their classifications by the transitive and intransitive RLS classifier are illustrated in Figure 3. From the results, we observe that the classifier using the transitive kernel can learn the relation to some extent, but the intransitive kernel is clearly better for this purpose.

5 Conclusion

In this paper the problem of learning intransitive reciprocal relations was tackled. To this end, we showed that existing approaches for preference learning typically exhibit strong stochastic transitivity as property, and we introduced an extension of the existing RankRLS framework to predict reciprocal relations that can violate weak stochastic transitivity. In this framework, the choice of kernel function defines the transition from transitive to intransitive models. By choosing a feature mapping based on the Kronecker-product, we are able to predict intransitive reciprocal relations. Experiments on benchmark problems in game theory and theoretical biology confirmed that our approach substantially outperforms the ranking approach when intransitive relations are present in the data. Given the absence of publicly available datasets on learning intransitive reciprocal relations, we are willing to share our data with other researchers, and in the future we hope to apply our algorithm in other domains as well.

Acknowledgments

This work was supported in part by the Academy of Finland. W.W. was supported by a research visit grant from the Research Foundation Flanders.

References

1. Diaz, S., Garcia-Lapresta, J., Montes, S.: Consistent models of transitivity for reciprocal preferences on a finite ordinal scale. *Information Sciences* **178**(13) (2008) 2832–2848
2. Azand, P.: The philosophy of intransitive preferences. *The Economic Journal* **103** (1993) 337–346
3. Tversky, A.: *Preference, Belief and Similarity*. MIT Press (1998)
4. Luce, R., Suppes, P.: Preference, utility and subjective probability. In Luce, R., Bush, R., Galanter, E., eds.: *Handbook of Mathematical Psychology*. Wiley (1965) 249–410
5. De Baets, B., De Meyer, H., De Schuymer, B., Jenei, S.: Cyclic evaluation of transitivity of reciprocal relations. *Social Choice and Welfare* **26** (2006) 217–238
6. Fisher, L.: *Rock, Paper, Scissors: Game Theory in Everyday Life*. Basic Books (2008)
7. De Schuymer, B., De Meyer, H., De Baets, B., Jenei, S.: On the cycle-transitivity of the dice model. *Theory and Decision* **54** (2003) 164–185
8. De Schuymer, B., De Meyer, H., De Baets, B.: Optimal strategies for equal-sum dice games. *Discrete Applied Mathematics* **154** (2006) 2565–2576
9. De Schuymer, B., De Meyer, H., De Baets, B.: Optimal strategies for symmetric matrix games with partitions. *Bulletin of the Belgian Mathematical Society - Simon Stevin* **16** (2009) 67–89
10. Makowski, M., Piotrowski, E.: Quantum cat’s dilemma: an example of intransitivity in quantum games. *Physics Letters A* **355** (2006) 250–254
11. Kerr, B., Riley, M., M.Feldman, Bohannan, B.: Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418** (2002) 171–174

12. Czárán, T., Hoekstra, R., Pagie, L.: Chemical warfare between microbes promotes biodiversity. *Proceedings of the National Academy of Sciences* **99**(2) (2002) 786–790
13. Kirkup, B., Riley, M.: Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissors in vivo. *Nature* **428** (2004) 412–414
14. Károlyi, G., Neufeld, Z., Scheuring, I.: Rock-scissors-paper game in a chaotic flow: The effect of dispersion on the cyclic competition of microorganisms. *Journal of Theoretical Biology* **236**(1) (2005) 12 – 20
15. Reichenbach, T., Mobilia, M., Frey, E.: Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* **448** (2007) 1046–1049
16. Boddy, L.: Interspecific combative interactions between wood-decaying basidiomycetes. *FEMS Microbiology Ecology* **31** (2000) 185–194
17. Sinervo, S., Lively, C.: The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **340** (1996) 240–246
18. Waite, T.: Intransitive preferences in hoarding gray jays (*perisoreus canadensis*). *Journal of Behavioural Ecology and Sociobiology* **50** (2001) 116–121
19. Pahikkala, T., Tsivtsivadze, E., Airola, A., Järvinen, J., Boberg, J.: An efficient algorithm for learning to rank from preference graphs. *Machine Learning* **75**(1) (2009) 129–165
20. Öztürk, M., Tsoukiàs, A., Vincke, P.: Preference modelling. In Figueira, J., Greco, S., Ehrgott, M., eds.: *Multiple Criteria Decision Analysis. State of the Art Surveys*. Springer-Verlag (2005) 27–71
21. Waegeman, W., De Baets, B., Boullart, L.: Kernel-based learning methods for preference aggregation. *4OR* **7** (2009) 169–189
22. Switalski, Z.: General transitivity conditions for fuzzy reciprocal preference matrices. *Fuzzy Sets and Systems* **137** (2003) 85–100
23. Hüllermeier, E., Fürnkranz, J., Cheng, W., Brinker, K.: Label ranking by learning pairwise preferences. *Artificial Intelligence* **172** (2008) 1897–1916
24. Fishburn, P.: *Utility Theory for Decision Making*. Wiley (1970)
25. Herbrich, R., Graepel, T., Obermayer, K.: Large margin rank boundaries for ordinal regression. In Smola, A., Bartlett, P., Schölkopf, B., Schuurmans, D., eds.: *Advances in Large Margin Classifiers*, MIT Press (2000) 115–132
26. Freund, Y., Yier, R., Schapire, R., Singer, Y.: An efficient boosting algorithm for combining preferences. *Journal of Machine Learning Research* **4** (2003) 933–969
27. Crammer, K., Singer, Y.: Pranking with ranking. In: *Proceedings of the Conference on Neural Information Processing Systems*, Vancouver, Canada. (2001) 641–647
28. Chu, W., Keerthi, S.: Support vector ordinal regression. *Neural Computation* **19**(3) (2007) 792–815
29. Hüllermeier, E., Fürnkranz, J.: Pairwise preference learning and ranking. In: *Proceedings of the European Conference on Machine Learning*, Dubrovnik, Croatia. (2003) 145–156
30. Chu, W., Ghahramani, Z.: Preference learning with gaussian processes. In: *Proceedings of the International Conference on Machine Learning*, Bonn, Germany. (2005) 137–144
31. Schölkopf, B., Smola, A.: *Learning with Kernels, Support Vector Machines, Regularisation, Optimization and Beyond*. The MIT Press (2002)
32. Shawe-Taylor, J., Cristianini, N.: *Kernel Methods for Pattern Analysis*. Cambridge University Press (2004)
33. Suykens, J., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J.: *Least Squares Support Vector Machines*. World Scientific Pub. Co., Singapore (2002)

34. Joachims, T.: Optimizing search engines using clickthrough data. In D.Hand, Keim, D., Ng, R., eds.: Proceedings of the 8th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'02), ACM Press (2002) 133–142
35. Magnus, J.R.: Linear Structures. Griffin, London (1988)
36. Weston, J., Schölkopf, B., Bousquet, O.: Joint kernel maps. In Cabestany, J., Prieto, A., Hernández, F.S., eds.: Computational Intelligence and Bioinspired Systems, 8th International Work-Conference on Artificial Neural Networks (IWANN 2005). Volume 3512 of Lecture Notes in Computer Science., Berlin Heidelberg, Germany, Springer-Verlag (2005) 176–191
37. Pahikkala, T., Pyysalo, S., Boberg, J., Järvinen, J., Salakoski, T.: Matrix representations, linear transformations, and kernels for disambiguation in natural language. *Machine Learning* **74**(2) (2009) 133–158
38. Pahikkala, T., Waegeman, W., Tsivtsivadze, E., Salakoski, T., De Baets, B.: Learning intransitive reciprocal relations with kernel methods. Submitted.
39. Frean, M.: Emergence of cyclic competitions in spatial ecosystems. In Whigham, P., ed.: SIRC 2006: Interactions and Spatial Process (Eighteenth Colloquium hosted by the Spatial Information Research Centre). (2006) 1–9